

(19)



Europäisches Patentamt  
European Patent Office  
Office européen des brevets

(11) Publication number:

**0 382 403  
A2**

(12)

## EUROPEAN PATENT APPLICATION

(21) Application number: 90301034.6

(51) Int. Cl.<sup>5</sup>: C12N 15/53, C12P 21/02

(22) Date of filing: 01.02.90

(30) Priority: 06.02.89 US 307366

(43) Date of publication of application:  
16.08.90 Bulletin 90/33

(84) Designated Contracting States:  
AT BE CH DE DK ES FR GB GR IT LI LU NL SE

(71) Applicant: UNIGENE LABORATORIES INC.  
110 Little Falls Road  
Fairfield, NJ 07006(US)

(72) Inventor: Betelsen, Arthur H.  
599 Maitland Avenue  
Teaneck, New Jersey(US)  
Inventor: Mehta, Nozer M.  
8 Apache Trail  
Rockaway, New Jersey(US)  
Inventor: Beaudry, Gary Agide  
81 Inwood Avenue  
Montclair, New Jersey(US)

(74) Representative: Perry, Robert Edward  
GILL JENNINGS & EVERY 53-64 Chancery  
Lane  
London WC2A 1HN(GB)

(54) Expression systems for amidating enzyme.

(57) Alpha-amidating enzyme is produced by recombinant DNA techniques recoverable in high yields and at high purity. Both eukaryotic and prokaryotic expression vectors are provided having a transcriptional promoter followed downstream by a DNA sequence which encodes amidating enzyme. The vector selected is one capable of directing the expression of polypeptides in the host selected, and preferred hosts are transected with the described vectors.

**EP 0 382 403 A2**

## EXPRESSION SYSTEMS FOR AMIDATING ENZYME

BACKGROUND OF THE INVENTION

This invention relates to the production of alpha-amidating enzymes through recombinant DNA techniques, and particularly to expression vectors and hosts capable of expressing alpha-amidating enzyme in high yields and at recoverable high purity.

The intracellular processing (cleavage and/or functional group modification) of precursor forms of native proteins following their translation from nucleic acid coding sequences has been clearly documented.

In general, mammalian cells and other eukaryotes can perform certain post-translational processing procedures, while prokaryotes cannot. Certain prokaryotes, such as *E. coli*, are widely employed as hosts for the production of mammalian proteins via recombinant DNA (rDNA) technology because they can be readily grown in batch fermentation procedures and because they are genetically well-characterized. However, many mammalian proteins require some type of post-translational processing, and if these proteins are produced by genetic engineering of *E. coli*, for example, the post-translational processing must often be accomplished by using complex, in vitro chemical procedures which are cost-prohibitive for large-scale production applications.

One type of processing activity involves the specific amidation of the carboxyl-terminal amino acid of a peptide or protein. Many naturally-occurring hormones and peptides contain such a modification, which is often essential if the protein is to be biologically active. An example is calcitonin, where the substitution of a non-amidated proline residue for the amidated proline of the native form results in a 3,000-fold reduction in biological activity. Other biological peptides requiring post-translational amidation for full activity include but are not limited to growth hormone releasing factor, other calcitonins, and calcitonin gene-related peptide.

The specific amidation of the carboxyl-terminal amino acid of a protein is catalyzed by alpha-amidating enzymes. The polypeptide sequences for many important biological proteins which require amidation for maximal efficacy, may be manufactured, for example, by genetic engineering techniques. However, the important and sometimes essential carboxyl terminal amidation must often be performed in vitro. It is desirable to avoid costly and cumbersome chemical amidation techniques at this point, and is therefore desirable to utilize an amidating enzyme to perform the specific amidation. However, alpha-amidating enzyme is not easily obtained in nature.

The presence of amidated peptides in a particular tissue is not necessarily synonymous with high levels of alpha-amidating enzyme. For example, rat anterior pituitary tissue contains high alpha-amidating activity but no known substrates [Eipper et al, PNAS 80, 5144-5148 (1983)]. Rat posterior pituitary tissue contains amidated peptides (oxytocin and vasopressin) but has very little alpha-amidating activity [Eipper et al., Endo 116, 2497-2504 (1985)]. Therefore, until individual tissues are tested for alpha-amidating activity, the presence or potential levels of the enzyme cannot be anticipated.

An even greater impediment to the availability of amidating enzyme obtained from natural sources is the usually low level of purity. Amidating enzymes obtainable from natural sources are contaminated with proteolytic enzymes and other impurities. Effective recovery of amidated product is greatly hindered when these impurity-laced enzymes are used to amidate a substrate comprised of L-amino acids. The presence of proteases, in particular, may break down the substrate and/or the amidated product and/or the amidating enzyme itself. Most biologically important polypeptides comprise L-amino acids, and are susceptible to this proteolytic breakdown and to other amidation-hindering impediments caused by impurities in amidating enzyme preparations.

Because nature provides few sources, low abundance and insufficient purity of alpha-amidating enzyme, there is a need for efficient methods of mass producing alpha-amidating enzyme recoverable in high yield and at high purity.

As used herein, the terms "amidating enzyme" and "alpha-amidating enzyme" refer to any agent capable of catalyzing the conversion of a peptidyl substrate to a corresponding peptidyl amide having an amino group in place of the C-terminal amino acid of said substrate.

BRIEF DESCRIPTION OF THE INVENTION

It is an object of the present invention to provide alpha-amidating enzyme recoverable in high yields



**TYPE A      $\alpha$ AE     COMPOSITE**

55

-353

5

10

15

20

25

30

35

40

45

50

55

60

65

70

75

80

85

90

95

100

105

110

115

120

125

130

135

140

145

150

155

160

165

170

175

180

185

190

195

200

205

210

215

220

225

230

235

240

245

250

255

260

265

270

275

280

285

290

295

300

305

310

315

320

325

330

335

340

345

350

355

360

365

370

375

380

385

390

395

400

405

410

415

420

425

430

435

440

445

450

455

460

465

470

475

480

485

490

495

500

505

510

515

520

525

530

535

540

545

550

555

560

565

570

575

580

585

590

595

600

605

610

615

620

625

630

635

640

645

650

655

660

665

670

675

680

685

690

695

700

705

710

715

720

725

730

735

740

745

750

755

760

765

770

775

780

785

790

795

800

805

810

815

820

825

830

835

840

845

850

855

860

865

870

875

880

885

890

895

900

905

910

915

920

925

930

935

940

945

950

955

960

965

970

975

980

985

990

995

1000

1005

1010

1015

1020

1025

1030

1035

1040

1045

1050

1055

1060

1065

1070

1075

1080

1085

1090

1095

1100

1105

1110

1115

1120

1125

1130

1135

1140

1145

1150

1155

1160

1165

1170

1175

1180

1185

1190

1195

1200

1205

1210

1215

1220

1225

1230

1235

1240

1245

1250

1255

1260

1265

1270

1275

1280

1285

1290

1295

1300

1305

1310

1315

1320

1325

1330

1335

1340

1345

1350

1355

1360

1365

1370

1375

1380

1385

1390

1395

1400

1405

1410

1415

1420

1425

1430

1435

1440

1445

1450

1455

1460

1465

1470

1475

1480

1485

1490

1495

1500

1505

1510

1515

1520

1525

1530

1535

1540

1545

1550

1555

1560

1565

1570

1575

1580

1585

1590

1595

1600

1605

1610

1615

1620

1625

1630

1635

1640

1645

1650

1655

1660

1665

1670

1675

1680

1685

1690

1695

1700

1705

1710

1715

1720

1725

1730

1735

1740

1745

1750

1755

1760

1765

1770

1775

1780

1785

1790

1795

1800

1805

1810

1815

1820

1825

1830

1835

1840

1845

1850

1855

1860

1865

1870

1875

1880

1885

1890

1895

1900

1905

1910

1915

1920

1925

1930

1935

1940

1945

1950

1955

1960

1965

1970

1975

1980

1985

1990

1995

2000

2005

2010

2015

2020

2025

2030

2035

2040

2045

2050

2055

2060

2065

2070

2075

2080

2085

2090

2095

2100

2105

2110

2115

2120

2125

2130

2135

2140

2145

2150

2155

2160

2165

2170

2175

2180

2185

2190

2195

2200

2205

2210

2215

2220

2225

2230

2235

2240

2245

2250

2255

2260

2265

2270

2275

2280

2285

2290

2295

2300

2305

2310

2315

2320

2325

2330

2335

2340

2345

2350

2355

2360

2365

2370

2375

2380

2385

2390

2395

2400

2405

2410

2415

2420

2425

2430

2435

2440

2445

2450

2455

2460

2465

2470

2475

2480

2485

2490

2495

2500

2505

2510

2515

2520

2525

2530

2535

2540

2545

2550

2555

2560

2565

2570

2575

2580

2585

2590

2595

2600

2605

2610

2615

2620

2625

2630

2635

2640

2645

2650

2655

2660

2665

2670

2675

2680

2685

2690

2695

2700

2705

2710

2715

2720

2725

2730

2735

2740

2745

2750

2755

2760

2765

2770

2775

2780

2785

2790

2795

2800

2805

2810

2815

2820

2825

2830

2835

2840

2845

2850

2855

2860

2865

2870

2875

2880

2885

2890

2895

2900

2905

2910

2915

2920

2925

2930

2935

2940

2945

2950

2955

2960

2965

2970

2975

2980

2985

2990

2995

3000

3005

3010

3015

3020

3025

3030

3035

3040

3045

3050

3055

3060

3065

3070

3075

3080

3085

3090

3095

3100

3105

3110

3115

3120

3125

3130

3135

3140

3145

3150

3155

3160

3165

3170

3175

3180

3185

3190

3195

3200

3205

3210

3215

3220

3225

3230

3235

3240

3245

3250

3255

3260

3265

3270

3275

3280

3285

3290

3295

3300

3305

3310

3315

3320

3325

3330

3335

3340

3345

3350

3355

3360

3365

3370

3375

3380

3385

3390

3395

3400

3405

3410

3415

3420

3425

3430

3435

3440

3445

3450

3455

3460

3465

3470

3475

3480

3485

3490

3495

3500

3505

3510

3515

3520

3525

3530

3535

3540

3545

3550

3555

3560

3565

3570

3575

3580

3585

3590

3595

3600

3605

3610

3615

3620

3625

3630

3635

3640

3645

3650

3655

3660

3665

3670

3675

3680

3685

3690

3695

3700

3705

3710

3715

3720

3725

3730

3735

3740

3745

3750

3755

3760

3765

3770

3775

3780

3785

3790

3795

3800

3805

3810

3815

3820

3825

3830

3835

3840

3845

3850

3855

3860

3865

3870

3875

3880

3885

3890

3895

3900

3905

3910

3915

3920

3925

3930

3935

3940

3945

3950

3955

3960

3965

3970

3975

3980

3985

3990

3995

4000

4005

4010

4015

4020

4025

4030

4035

4040

4045

4050

4055

4060

4065

4070

4075

4080

4085

4090

4095

4100

4105

4110

4115

4120

4125

4130

4135

4140

4145

4150

4155

4160

4165

4170

4175

4180

4185

4190

4195

4200

4205

4210

4215

4220

4225

4230

4235

4240

4245

4250

4255

4260

4265

4270

4275

4280

4285

4290

4295

4300

4305

4310

4315

4320

4325

4330

4335

4340

4345

4350

4355

4360

4365

4370

4375

4380

4385

4390

4395

4400

4405

4410

4415

4420

4425

4430

4435

4440

4445

4450

4455

4460

4465

4470

4475

4480

4485

4490

4495

4500

4505

4510

4515

4520

4525

4530

4535

4540

4545

4550

4555

4560

4565

4570

4575

4580

4585

4590

4595

4600

4605

4610

4615

4620

4625

4630

4635

4640

4645

4650

4655

4660

4665

4670

4675

4680

4685

4690

4695

4700

4705

4710

4715

4720

4725

4730

4735

4740

4745

4750

4755

4760

4765

4770

4775

4780

4785

4790

4795

4800

4805

4810

4815

4820

4825

4830

4835

4840

4845

4850

4855

4860

4865

4870

4875

4880

4885

4890

4895

4900

4905

4910

4915

4920

4925

4930

4935

4940

4945

4950

4955

4960

4965

4970

4975

4980

4985

4990

4995

5000

5005

5010

5015

5020

5025

5030

5035

5040

5045

5050

5055

5060

5065

5070

5075

5080

5085

5090

5095

5100

5105

5110

5115

5120

5125

5130

5135

5140

5145

5150

5155

5160

5165

5170

5175

5180

5185

5190

5195

5200

5205

5210

5215

5220

5225

5230

5235

5240

5245

5250

5255

5260

5265

5270

5275

5280

5285

5290

5295

5300

5305

5310

5315

5320

5325

5330

5335

5340

5345

5350

5355

5360

5365

5370

5375

5380

5385

5390

5395

5400

5405

5410

5415

5420

5425

5430

5435

5440

5445

5450

5455

5460

5465

5470

5475

5480

5485

5490

5495

5500

5505

5510

5515

5520

5525

5530

5535

5540

5545

5550

5555

5560

5565

5570

5575

5580

5585

5590

5595

5600

5605

5610

5615

5620

5625

5630

5635

5640

5645

5650

5655

5660

5665

5670

5675

5680

5685

5690

5695

5700

5705

5710

5715

5720

5725

5730

5735

5740

5745

5750

5755

5760

5765

5770

5775

5780

5785

5790

5795

5800

5805

5810

5815

5820

5825

5830

5835

5840

5845

5850

5855

5860

5865

5870

5875

5880

5885

5890

5895

5900

5905

5910

5915

5920

5925

5930

5935

5940

5945

5950

5955

5960

5965

5970

5975

5980

5985

5990

5995

6000

6005

6010

6015

6020

6025

6030

6035

6040

6045

6050

6055

6060

6065

6070

6075

6080

6085

6090

6095

6100

6105

6110

6115

6120

6125

6130

6135

6140

6145

6150

6155

6160

6165

6170

6175

6180

6185

6190

6195

6200

6205

6210

6215

6220

6225

6230

6235

6240

6245

6250

6255

6260

62

CAGGAGAAAGTATTAGACCAGGCGTATTTCTATTCAGCTCTTCCAGCTCTAGGAGAAAGCGAAGATGTTCAATGCCACAAGTATAATCTACAGAA 1253  
 . . . . . (419)  
 5 AGACAGAACTCGGTTCAGACCTCTGCTGACATTCGAAACGCTGCTCCAGAAAGACACCTTCGTCGGTCTCAGCCACAGAAAGCTCCAGACGATGACCA 1353  
 . . . . .  
 ATGGCGTAATGCTATCTAGTCAAGACAGGATCCAGATTCACACAGCTAGAGTCACTCTGAGCCAGCTGAGACAGACGTTTCTGTTCCAGCAO 1453  
 . . . . .  
 10 CTTGCGGAAGCGCTTGGCAACCAAGCCCTCAGGAGATTTCCATGTGGAAGAACTGACCTGCGCTGAGTCTACTTGTACCAGGCGAGCTTCTG 1553  
 . . . . . (519)  
 CGGTGCGCTGAGTCTAGGATAAAGCTAGTGAATTTCCAGAGCGTGAACATGTTGCGATGGAAGCTTTTACAGCAAGTTTGTATCCAGCAAG 1653  
 . . . . .  
 15 AGGTCTTGGCCCAATTGAAGACACACCATCTGCTGATTCAGCCAAATATGCTGAATCTCCAGTCCAGTCCAGCAAGCTGTTTATTTACACAC 1753  
 . . . . .  
 CGCTTGAACATAGATAGATGCAAAATATTGCGTCAACATGCTGCTTCCAGCGTCTCAAAATGGACCCGATAGCAAGAGCGGCTCTCTTAA 1853  
 . . . . . (619)  
 TTCTGCGAAGGACCATGCAAGCTGCGAGTGACCAAAATCATTCTGCGACCCACCCATGCTGCTGCGAGCCAGTACTGGAGCTGCTTCTGTTGAGA 1953  
 . . . . .  
 20 CGGTACTGTAAAGTCTGCGATTTGCGATTTTCAACAGCGGAAGTCTGCTGCGAGTCCGAGGAGAGTCTCTGCAAGCACTCTACCGCAGGCAO 2053  
 . . . . .  
 TTCAGTGTCTGACAGTTTGGCTTGTGCTTCAATGGACCACTTGTGCTGCGAGAGGCAAAATGCGCAATCCAATGCTTCAAAAGTACACCA 2153  
 . . . . . (719)  
 AAGAATTTGCGAGAGATTAAGCAGGATCATTGGAAGCAATGCTTTGCCATTTCATATATACAGGTTTCTCTTCCGCTAAACGGGAAGCCCTTA 2253  
 . . . . .  
 25 CTTTGGAGACCAAGACCGCTGCAAGCATTGTCATGAATTTTCCAGTCCGGAATATAGACGCTTCAAGCCAGTACCGCAAGCACTTCGACATGCT 2353  
 . . . . .  
 CATCATATTTGCTGCTCTGAGATGCGACTGTGTACATTGGAGACCCACACACACACCGTGTGGAAGTTCACCTGACTGAAAAAATCCAGCATCGT 2453  
 . . . . . (819)  
 30 CAGTCAAAAAGCGCTGCCATTGAAGTCCAGGAAATCAAGCCGAGGCACTGTTGAACCCAAAGTGGAGACAAAGCTACCTCTCAGAAATGCAAGAT 2553  
 . . . . .  
 GCAAGAGAAACAGAAAGTACGACAGAGCGCGCTCGGAGTCTCGTCTGCTCATTACAACCTTCTGCTTATCTGCTGCTGCTGCTGCTGCTGCTGCT 2653  
 . . . . .  
 35 GTCATGTTTATTCGCTGGAATAATCAAGCGCTTTGCGACGATCATGACCCCAAGCTCGAGTCAAGTCTGGAAGAGCTCTGGAAGAGAGGAGGAG 2753  
 . . . . . (919)  
 GACGAGAAAG 2853  
 . . . . . (944)  
 AGCGCTCTGAGAAATAG 2943  
 . . . . .

As used herein, the term "stringent conditions" means 2 x SSC (0.3M sodium chloride and 0.03M sodium citrate) at 62°C.

The present invention also provides expression vectors for directing expression of alpha-amidating enzyme in both prokaryotic and eukaryotic systems. For example, an expression vector is provided which is capable of directing, in a prokaryotic host, the expression of a polypeptide sequence of an alpha amidating enzyme, said vector comprising a transcriptional promoter followed downstream by a first DNA sequence having an amidating enzyme-coding region, said first sequence being sufficiently homologous to a natural DNA sequence for expressing natural amidating enzyme to undergo hybridization with said natural sequence under stringent conditions, and said first sequence including an initiating methionine codon within about 50 nucleotides of the start of said enzyme-coding region.

Likewise, an expression vector is provided, in another embodiment of the invention, which is capable of directing the expression of a polypeptide sequence of an alpha-amidating enzyme in a eukaryotic host, said vector comprising a transcriptional promoter followed downstream by a first DNA sequence having an amidating enzyme-coding region, said first sequence being sufficiently homologous to a natural DNA sequence for expressing natural amidating enzyme to undergo hybridization with said natural sequence under stringent conditions, and said first sequence including a stop codon upstream from a sequence which would otherwise code for a membrane spanning domain.

This first sequence should be followed by a sequence specifying the addition of poly A to the

messenger RNA generated by transcription from said promoter.

As used herein, the term "membrane spanning domain" is a DNA sequence which, as determined by the test of Kyte & Doolittle, J. Mol. Biol., Vol. 157, pp. 105-132 (1982) (the entire disclosure of which is hereby incorporated by reference), codes for an amino acid sequence of sufficient hydrophobicity, length, structural character, and the like to become fixed in the membrane. For example, this may occur as a protein is synthesized on a membrane-bound ribosome or, alternatively, the amino acid sequence coded by the membrane spanning domain may become associated with other areas of the protein of which it is a part, such that the sequence becomes inserted into the hydrophobic environment of the membrane post-translationally. Membrane-spanning domains are discussed in more detail in Von Heijne, Sequence Analysis in Molecular Biology: Treasure Trove or Trivial Pursuit, pp. 81-121 (Acad. Press 1987), the teachings of which are hereby incorporated by reference.

The base numbers utilized herein are the numbers specifically stated for any DNA sequence expressly set forth together with base number references. For all sequences for which base numbers are not expressly assigned herein, the bases shall be consecutively numbered with base number 1 being the first base of the first codon that is expressed by the sequence being discussed, and the amino acid numbers are consecutively numbered with the first being the amino acid expressed by bases 1-3.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a flow chart for construction of a mammalian expression vector for alpha-amidating enzyme.

Figure 2 is a flow chart for construction of a prokaryotic expression vector for alpha-amidating enzyme.

Figure 3 is Coomassie Blue-stained SDS-PAGE electrophoretogram of the insoluble protein fraction from *E. Coli* JM105 carrying the plasmids indicated (which have the characteristics set forth in Example 1) when cultured with (+) or without (-) IPTG added to the growth medium (C=insoluble proteins of *E. coli* JM105 carrying pKK233-2).

Figure 4 is a Western Blot of the gel shown in Figure 3 wherein, following protein transfer to nitrocellulose, the filter was treated with rabbit anti AE antisera and the alkaline phosphatase-conjugated anti rabbit Ig, followed by chromogenic substrate for the alkaline phosphatase.

#### DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

In accordance with the invention, expression vectors suitable for prokaryotic systems and expression vectors suitable for eukaryotic systems are prepared. DNA encoding amidating enzyme useful in these vectors may be isolated as taught in EPC application No. 88307533.5 and published as Publication No. 0308067 the entire disclosure of which is incorporated herein by reference. Alpha-amidating enzymes have been isolated from rat medullary thyroid carcinoma tissue or conditioned media from a rat cell line, and purified to homogeneity as taught in the above-identified parent application and in grandparent application Serial Number 655,366 filed September 27, 1984, now issued as U.S. Patent 4,708,934, priority of which has been claimed and the entire disclosure of which has been incorporated herein by reference. Amino acid sequences have been determined for the purified alpha-amidating enzyme, and these sequences have been used to project a variety of oligonucleotide probes which have been radiolabelled and utilized for isolating cDNAs for amidating enzyme.

The isolated cDNAs have been used to screen libraries prepared, for example, from the total RNA of rat medullary thyroid carcinoma tissues, their derived cell lines or from cell lines known to produce amidating enzyme, for example, biological deposit IVI 10028 (In Vitro International, Linthicum, Maryland) (Rat MTC tissue) or derived cell line IVI 10029. Total RNA was prepared and Poly-A RNA was selected with oligo DT cellulose. cDNAs were prepared by well known methods utilizing first reverse transcriptase and then a DNA polymerase. The cDNA was used to generate cDNA libraries in the vector  $\lambda$ gt 11 and the recombinant DNA's were packaged *in vitro* to form infectious bacteriophage particles.

Extracts for packaging are commercially available for example from Promega Biotech or Clontech Laboratories or can be prepared according to methods well-known in the art. The phage were screened with radiolabelled oligonucleotide probes prepared as set forth above. Screening for bacteriophage containing alpha amidating enzyme cDNA ("AE cDNA") was accomplished by plating samples of bacteriophage and lifting the phage onto nitrocellulose filter discs. Hybridization with two or more radiolabelled AE-specific

Oligonucleotide probes denoted AE4, AE5, AE8 and AE9 at pages 61-64 of EPC application No. 88307533.5 and published as Publication No. 0308067 are especially preferred when screening libraries prepared from biological deposit IVI 10029 identified above.

10 of lines indicate amino acid numbers:

## TYPE A



5 TCATGTATTACATCGAAGCCAAATATGCACTTCTCTCATGACCTGTACAAAGAACCTGGCTCCAGATATGTTCAAACTATCCACGACGCCAATAT 1055  
 y y m e n h y a i s i t c t c h e v a p d a f r t i p a o o n i  
 10 CCCAATTCCTGTCAAACCCGACATCGTTATGATCCAGCGGCATCACAAGAACCCAGAAACAAAGAAAGACTGCTTAAATGCAGCAGCCAAACACGGA 1155  
 p i p v r p d e v n h g h b h e o o h o h a i n q q p h q e  
 15 GAGCAAGAGTATTAGACGAGATTTCCATGTGGAAGAACTGGACTGGCTGGAGTGTACTTGTACCAGCCAGGTTTCTGGGCTGGCCCTCGATT 1255  
 o e v l o q d i b v e o o i d w p s v y l i p s q v a s v o i d a (A19)  
 20 CTAAGAATAACCTACTGATTTTCCACACAGCTGACCATGTTTCCGATCGAAACTCTTTTACAGCAAGTTCCTTACCAGCAAGAGGCTTGGCCCAAT 1355  
 l n n i v i l h r g d h v w d g n o f d h i v y q q r e i p i  
 25 TGAAGAAGACACCATCTCGTCATTGACCCAAATATGCTGAAATCTCCAGTCCAGTGGCAAGAACCTGTTTATTACCACAGCGCTTGACCATAGAT 1455  
 o o d t l i l v i d p n n o i l q o s g h n i l y l p h s l s i d  
 30 ACAGATCGAAATATTGGTCAAGATGTGGCTCTCCACAGCTGTTCAAATTCGACCCGATACCAAGCAAGCCCTCTCTTAATTCGCAAGGAGCA 1555  
 t d s a y w t d v a i h q v t h i d p h s h o s p i l i l g r o (519)  
 35 TCCAAGCTGGGAGTGACCAAAACCATTTCTCCAGCCCAAGGATGTGGCTGTGGAGCCAGTACTGGAGCTGTCTGCTGTGACAGCGTTACTGTAAACAG 1655  
 q p g s d q n h t c q p t d v a v o p s t s a v f v d g y c n o  
 40 TCGGATTGTCAGATTTTACCAAGCCGAAAGTTCTGACCCAGTGGGAGAGAGCTCTTGGCAAGCAGTCTAGCCGACCCAGTTCAGTGTCTCTCAC 1755  
 r i v q l s p g h i v t q u s o s s s s p r p s q f a v p h  
 45 AGTTTGGCCCTTGTGGCTCATTTGGACAGTTGTGTGTGGCAGACAGGAAATGGCCGAAATCCATGCTTCAAACGACACCAAGAATTTGTGAGAG 1855  
 a l o l v p h i d q i c v a d r o n g r i q c i h t d t h o f v r o (619)  
 50 AGATTAGCAGCCATCATTTGCAAGCAATGCTTTGCCATTTTATATATACCAAGTTCCTCTTGGCCGTAACCCGAGCCCTTACTTTGACAGACCAAGA 1955  
 l k h o s f s r v f a i s y i p s i l l a v n g h p y f s d q o  
 55 GCGCTGCAAGGATTTGTGATGAATTTCCAGTGGGAAATATAGACCTCTTCAAGCCAGTACCAAGCAGCTTGGACATGCTCATGATATTGTGGCT 2055  
 p v q s i v n n i s s o l i d v i h p v r h h f d o p h d i v o  
 60 TCTGAAGATCGGAGTGTGTACATTGCGACGCCACACACAAACCCGTGTGCAAGTTCACCTGACTGAAAAATGAGCATCGGTCAATTAAGAGGCTG 2155  
 s o d s t v y i s d a b t n t v w h f t i t o h n o h r o v h h a s (719)  
 65 GCATTGAAGTCCAGGAAATCAAAGAAGCCGAGTGTGTAACCCAAAGTGGAGAACAAACCCCTCTCAGAAATTCAGAGATGCAAGAGAAACA 2255  
 l e v q o i k o a o a v v o p h v o n h p t u s o i q h u q o h q  
 70 GAAACTGACACAGAGCCCGCTCGGAGTGTCCGTGTTTCATACCAACCTTCTGGTATTCCTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2355  
 k i s t o p s s g v o v v i l t t i l v i p v i v l i o i v o i l  
 75 CCGTGGAAAAATCAAGCCCTTTGAGGAAAGCAACCCGCTTAAATTTGGAAATTTCTTCCAAATCGAAAAATCTACAGCAGAAAGGCTTTG 2455  
 r n h h s r a f s s h s s s i n i s n f l a s r h s y o r h s f d (819)  
 80 ACCGAGTACGACAGAGCCGAGTACCAAGAGAAAGATGAGCAGCAGGGAAGTGAAGAGGAGTACTCGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCT 2555  
 r v t t o s s d q o h d o d d s s s o p i p h p a p o  
 85 CTCTGAGCTCCAGCTTCCCGGAGTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2655  
 s (853)  
 90 GTAACTGTACTGTCAGTGTGGAGTGTACACCTTTATTACTTATTTGCTTCCCTTCCCTTCTGCTTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 2755  
 CTCAGTGGCATTGTCTTATATGAAGTTAGCGTAGAGACCGCCCTTCTCTTCCAGGTACCGTACCCGAGGAAGTTTACTCATTACATTGAGAC 2855  
 GTTACTTCTGAGTGTAAATAGCCCTATTCTGCTTGAACACAGTATCTCCAGTCTCCAGCAGTCTCTTCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 2955  
 AGCATTCTCAGCTGTGGAGTGAAGAGAACCACTGCCACAGCAGGAAAGCTGCTAATCTCTCTATTTTAAATCACTAACATTATATGCA 3055  
 ATCAGAGAAATTTAAAGCTCTCTATTTAAATTTCTTTTAAATTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3155  
 GTTACACACTATTTTCAAACTCTGAATGTAATTTGTGTAATAAGTGTTCAGACCAATTACCAAAAAA  
 3214

Another type of cDNA sequence isolated by the oligonucleotide hybridization screening set forth above is represented by Diagram B below where the numbering and other conventions are the same as those stated for Type A above:

[illegible]

11

preferred that the gene sequences that encode the natural secretory signal sequence be removed, that an initiation codon be placed within about 50 nucleotides of the gene sequences encoding the start of the mature protein corresponding to alpha-amidating enzyme, and that the gene sequences encoding the membrane spanning domain in the C-terminal region not be translated. The initiating codon is of course in-frame with the sequence which encodes the enzyme and, in some preferred embodiments, is upstream from the region, sometimes immediately upstream.

When the AE cDNA was expressed in *E. Coli*, it was discovered that the natural gene sequence contained a cryptic *E. Coli* ribosome binding site ("RBS") and initiation codon internal to the natural initiation sequences. This resulted in the production of an N-terminally truncated amidating enzyme protein. While this did not prevent the production of the desired product in *E. Coli*, the coexistence of the correctly initiated and internally initiated products complicates the processing and purification of the recombinant product to a useful form and is therefore undesirable. To eliminate the unexpected, undesired product, it was necessary to eliminate either the ribosome binding site, the internal initiation codon or both of these.

For example, in certain preferred embodiments of the invention, a valine codon which, in prokaryotic systems, codes an initiating methionine, is altered by a point mutation to an equivalent non-initiating valine codon at bases 661-663 (of the cDNAs of either diagram A or B). In lieu of this point mutation or in addition thereto, applicants, in other preferred embodiments, delete or substantially modify any region coding for a ribosome binding site which occurs just upstream of an internal initiation site, and more preferably any ribosome binding site whenever one may occur. These modifications are made to substantially eliminate internal initiation such that the protein expressed because of internal initiation is not observed as a separate band following electrophoresis.

To obtain expression of secreted, active alpha amidating enzyme protein, from a recombinant eukaryotic host cell line it was necessary to remove the gene sequences encoding the transmembrane domain found in the C-terminal region of the natural gene sequences. For the type A cDNA this has been done by truncation of the protein coding region through introduction of a stop codon at or near to where the natural amidating enzyme is post-translationally processed in some natural systems as explained in detail below. For the type B cDNA this has also been done by introducing a new stop codon in the region of the enzyme protein where the natural type B amidating enzyme is post-translationally processed (see below). This should not be taken to exclude the possibility that in some host cell systems it may be preferable to express the entire naturally occurring gene sequences. Because the type B cDNA contains sequences with the characteristics of an unprocessed intron there may be a difficulty in expressing this cDNA in some eukaryotic host cells. These cells may not efficiently produce an mRNA from the type B gene due to the presence of the paired splice donor and acceptor sites. Elimination of the acceptor site might therefore be necessary to allow for efficient expression of type B AE cDNA.

We have discovered that the carboxyl end of the naturally occurring 75 kD alpha amidating enzyme protein occurs beyond amino acid position 709 (814 of type B). To produce the 75 kD protein (87 kD of type B) in a recombinant DNA host cell, a stop codon has been introduced into the cDNA by mutation of the codon for the lysine of amino acid position 716 (821 of type B). This modification has been made using oligonucleotide directed site specific mutagenesis. Such mutagenesis can be accomplished in a variety of ways. The methods have been reviewed extensively in the molecular biology literature. The general method that we have used was described by Taylor, J.W. et al. (1985), *Nucl. Acids Res.*, 13: 8749-8764; Taylor, J.W. et al. (1985), *Nucl. Acids Res.*, 13: 8764-8785; Nakamaye, K. and Eckstein, F. (1986), *Nucl. Acids Res.*, 14: 9679-9698. The reagents needed to practice this method are available in the form of a mutagenesis kit from Amersham Corporation.

The mutation of the sequence that we have produced changes the AAA lysine codon to a TAA stop codon. The oligonucleotide used for the mutagenesis incorporated this change but was otherwise identical in sequence to the naturally occurring cDNA sequence for the respective enzyme (type A or type B) being mutated.

We have also discovered that a naturally occurring shortened form of the alpha amidating enzyme protein is produced by processing of the type B protein at the internal region of the protein that is unique to the type B enzyme protein. This results in an enzyme product that is approximately 43 kD in molecular mass. Without intending to be bound by theory, it is believed that the DNA sequence upstream from the intron region is sufficient to code for a polypeptide capable of exhibiting significant alpha-amidating activity. Accordingly polypeptides which are easy to recover and which are capable of expressing alpha-amidating activity may be encoded by cDNAs which are significantly truncated by placement of a stop codon somewhere in the intron region of type B cDNA in just before or after the corresponding location where this intron is missing from TYPE A cDNA. Preferred truncation results from placement of a stop codon within about 30 bases of the beginning of the of the intron region, preferably immediately downstream therefrom.

To enable the production of one preferred short form of alpha amidating enzyme protein in recombinant host cells, a modified cDNA is created having a stop codon in place of the lysine codon at amino acid position 436 of the type B cDNA. This mutation was accomplished by oligonucleotide directed site specific mutagenesis of the type B AE cDNA.

5 While the shortening of the amidating enzyme protein by introduction of the stop codon at amino acid position 436 of the type B cDNA gives a protein that most closely approximates the one produced naturally by proteolytic cleavage of the primary translation product (or some other cleavage intermediates in the biosynthetic pathway), a further shortening of the amidating enzyme protein may also result in production of an active product in recombinant DNA host cells. We have modified the AE cDNA in several other ways to  
10 create such shorter forms of protein. In one example, we have used oligonucleotide directed site-specific mutagenesis to convert a tyrosine codon at amino acid position 396 of the type B cDNA to a stop codon. This change will result in a protein that is approximately 39 kD when the cDNA is translated and processed. In a second case, we have utilized the naturally occurring Bam H1 enzyme recognition site of the type B cDNA to introduce a stop codon by linker mutagenesis. This method is well known in molecular biology and  
15 simply involves the cleavage of the cDNA followed by ligation to a double stranded synthetic linker fragment that is complimentary to one end of the cleaved cDNA and that introduces an in frame stop codon just beyond the cleavage site. We have used an oligonucleotide fragment with the following sequence to accomplish this modification:

20



25 This linker introduces a stop codon following the histidine codon at amino acid 469. Translation and processing of the cDNA once it has been modified in this fashion results in the synthesis of a protein of approximately 46 kD.

Preferred placement of a truncating stop codon is within about 15 bases of a DNA sequence which codes for consecutive basic residues (usually a Lys-Lys) and especially immediately upstream therefrom.  
30 Without intending to be bound by theory, it is believed that the natural polypeptide coded by the cDNAs of type B is processed, during post-translational modifications which occur during natural expression of amidating enzyme, at or near such consecutive basic residues, for example, the consecutive lysines coded within the intron region of the cDNA of diagram B. Even when the inserted stop codons are not intended to truncate the expressed polypeptide in the above-described manner, it is preferred that the inserted stop  
35 codon be placed within about 20 bases, and preferably immediately upstream from, DNA sequences coding for consecutive basic amino acid residues. Insertion of stop codons at these positions will likely result in expression of a polypeptide resembling certain natural amidating enzymes after they have undergone post-translational processing.

For cytoplasmic expression in prokaryotic systems, any signal sequence coding regions (for example,  
40 the first bases of both the type A and type B cDNAs diagrammed previously) are preferably eliminated and a methionine initiator codon is inserted within about 50 nucleotides of the beginning of the region which codes for amidating enzymes.

An alternative embodiment for prokaryotic expression eliminates any coding sequences for signal sequence or proenzyme sequence and inserts an initiator methionine codon within about 50 nucleotides of  
45 the beginning of the region which codes for amidating enzyme. In many natural AE cDNAs, this corresponds to the beginning of the region which encodes ser-x-ser (X being phe or leu). See, for example, bases 124 to 132 of the sequence for type A or type B cDNA. In some embodiments secretion of alpha amidating enzyme may be desirable. In this case it is preferable to retain the signal sequence coding regions, or alternatively to replace them with heterologous sequences that can serve the same function, for  
50 example, the signal sequences of the bacterial OMP A protein.

It will be readily apparent to those skilled in the art that numerous mutations and truncations of the DNA sequences set forth herein for encoding amidating enzyme are possible within the scope of the invention and that such modified sequences would code for polypeptides capable of functioning as amidating  
55 enzymes. Accordingly, applicants' claims should be construed to include all functional equivalents of DNA sequences, expression vectors and host cells specifically set forth.

Examples of prokaryotic expression vectors which may desirably be modified to include DNA sequences encoding amidating enzyme in accordance with the invention include but are not limited to pKK233-, pKK322-2, pPROK-1, pkT279,280,287, pPL lambda, pYEJ001, pKC30, pPROK-C, all commercially

available. Prokaryotic hosts which may be transfected with expression vectors in accordance with the invention include but are not limited to C600, LE392, RR1, DH1, SF8, all commercially available.

Eukaryotic expression vectors which may desirably be modified to include DNA sequences encoding amidating enzyme in accordance with the invention include but are not limited to pMAMNeo, pdBPVMMTNeo, pRSV, peuKC1, pCH110, all commercially available. Appropriate yeast vectors may also be used. Preferred eukaryotic hosts may be transfected with expression vectors in accordance with the invention include but are not limited to IVI deposit 10029, HeLa, CV1, C127, CHO (Chinese Hamster Ovary) and COS.

# EXAMPLE 1

## Expression of Alpha Amidating Enzyme Proteins in E. Coli

In order to express alpha amidating enzyme in E. Coli (see the flow chart of Figure 2), a cDNA fragment having the sequence set forth in diagram A, above, was digested With KpnI and Hind III and the fragment of about 2.1 kb was isolated. To build back a amino terminus corresponding to a natural mature enzyme, an oligonucleotide linker with the sequence

5' CATGTCATTTTCCAATGAATGCCTTGGTAC 3'  
3' AGTAAAAGGTTACTTACGGAAC 5'

was ligated to this DNA fragment. The resulting fragment contained one Nco I compatible sticky-end and one Hind III sticky end. The E. Coli expression vector pKK233-2 was obtained commercially from Pharmacia and digested with restriction enzymes Nco I and Hind III. The large linear fragment was isolated and ligated to the linker adapted cDNA fragment. The ligation mixture was used to transform competent E. Coli JM105. Transformants were selected by ampicillin resistance and the clones isolated were analyzed for the recombinant plasmid by restriction enzyme and DNA sequence analysis to confirm the structure of the expression vector (hereafter "pAE12") that they contained. The expression vector contains the hybrid trp-lac promoter which is repressed by the lac repressor and inducible by treatment of the cells with isopropyl-thiogalactoside (IPTG). Upstream from the initiator methionine the vector also contains the sequences of a strong ribosome binding site.

To obtain expression of the alpha amidating enzyme in the E. Coli, the recombinant cells were grown with shaking in LB-broth at 37° C to an OD<sub>600</sub> of 0.4. IPTG was added to the culture to a final concentration of 1mM and the growth was allowed to continue at 37° C with shaking for three to five hours. Cells were collected by centrifugation of the culture and the supernatant was discarded. The cells were resuspended in buffer containing a cocktail of protease inhibitors, treated with lysozyme and then sonicated to lyse the cell membranes. The lysates were centrifuged at 12,000 x g to separate the soluble and insoluble fractions of the cells. Each fraction was analyzed by SDS-PAGE and protein staining. The alpha amidating enzyme protein was readily identified as an IPTG inducible product in the insoluble protein fraction. Since the initial expression plasmid did not contain a stop codon specified by the alpha amidating enzyme gene sequences, the inducible product formed contains sequences specified by downstream vector DNA fused to the C-terminal of the alpha amidating enzyme protein sequences. In addition, the induced insoluble protein also contained a smaller amidation enzyme specific protein that represented a product formed by internal initiation of protein synthesis at a cryptic RBS and initiation codon (amino acid position 221 of the alpha amidating enzyme sequence).

To remove the unwanted sequences from the C-terminal portion of the expressed product, a mutation of the lysine codon at position 716 of the type A sequence was made to generate a TAA stop codon at this position. The mutated cDNA was then digested with Kpn I and Eco R1 and used to replace the original Kpn I-Eco R1 fragment in the initial expression vector pAE12. In a similar fashion, the type B cDNA sequences were mutated at the comparable position (amino acid 821) to create a stop codon and the Kpn I-Eco R1 fragment from the mutated type B cDNA was used to replace the corresponding fragment in pAE12. The two expression plasmids so created pAE24 (type A) and pAE25 (Type B) were then used to transform

JM105. The resulting strains were cultured for expression as was done previously for pAE12-containing strains. The pAE24 was found to produce two IPTG inducible, insoluble proteins of approximately 75 kD and 55 kD while the pAE25 was found to produce two IPTG inducible insoluble proteins of about 87 kD and 67 kD. Again, the small protein in each of these pairs represents the unwanted amino-terminally truncated product from either the type A or type B cDNA.

To eliminate the initiation of protein synthesis at the cryptic internal ribosome binding site and initiation codon (amino acid position 221) the GTG start codon, (GTG can serve as an initiator met codon in bacteria), was converted to a GTT codon that cannot initiate protein synthesis but which still encodes the valine that is normally found at this position in alpha amidating enzyme proteins encoded by natural genes. When the mutated region of the cDNA was substituted for the natural sequence in the expression vectors pAE24 and pAE25, two new vectors were created, pAE31 and pAE32. Transforming *E. coli* JM105 with these modified expression vectors and testing protein production from the resulting recombinant strains indicated that this mutagenesis was effective in eliminating the unwanted internal initiation. The IPTG induced product from the host cells carrying pAE31 was found to be 75 kD while that from cells transformed with pAE 32 was found to be 87 kD.

Since we have found that naturally occurring amidating enzyme from type B cDNA is post-translationally processed to give proteins of approximately 43 kD, we have prepared a series of mutations in type B AE cDNA that allows expression of proteins that terminate at or near the position where the naturally processed enzyme ends. Two of these mutations were prepared by oligonucleotide mutagenesis while a third was created by adapter-linker mutation as indicated above. When cDNAs carrying these mutations were used to replace the corresponding segments of pAE32, transformed into JM105 and analyzed for protein production in experiments similar to those described above, truncated alpha amidating enzyme proteins were detected. With a mutation at amino acid position 396 of type B cDNA changing a natural tyrosine codon to a stop codon (pAE36), a 39 kD enzyme protein was found while a linker mutagenesis that ended translation at the histidine codon of amino acid 464 resulted in a vector, pAE51, which produced a recombinant alpha amidating enzyme protein of 46 kD following transformation and induction of *E. coli* JM105.

All of recombinant alpha amidating enzyme proteins produced in *E. coli* described above were found to segregate with the insoluble fraction of the cell extracts. The enzymes could be rendered soluble and active by treatment with 8 M urea followed by rapid dilution in 50mM Tris-HCl pH7. When *E. coli* JM105 carrying pAE12 was grown and induced with IPTG as described, the alpha amidating enzyme proteins were present at levels of at least 30 mgs per liter of bacterial culture.

Representative samples of the induced insoluble protein produced in *E. coli* carrying AE expression plasmids are shown in Figures 3 and 4.

## EXAMPLE 2

### Generation of mammalian expression vector pd BPV-MMTNEO-AE<sub>A75</sub>

To generate a mammalian expression vector which expresses and constitutively secretes 75 kD type A alpha amidating enzyme from mammalian cells (see the flow chart of Figure 1), the following was performed:

1) The intermediate expression vector pdMMTNeo (commercially available from American Type Culture Collection) (as shown) was digested with Bgl II. The linear form was isolated and purified.

2) The recombinant type A cDNA containing the full prepro sequence and an artificial stop codon TAA at position 2146-2148 was isolated by sequential digestion with Bgl I and Xho I. The fragment corresponding to alpha amidating enzyme was then isolated and purified.

3) The insert (type A alpha amidating enzyme) and vector (pdMMTNeo) were mixed and the corresponding ends were made flush using the Klenow fragment of DNA polymerase I. The 5' protruding segments were filled in with added dNTP, and the 3' protruding segments were digested back to produce a flush end (alternatively sequential S1 nuclease and Klenow + dNTP could be utilized for producing flush ends). The flush ended molecules were then ligated for 16 hours at 15°C.

4) The ligated material was then transformed into *E. coli* RRI. Recombinant clones were selected in the presence of 50 ug/ml ampicillin. The orientation of the insert in the recombinant clones was verified using a battery of restriction enzymes. One clone which was referred to as pdMMTNeo  $\alpha$  -AE<sub>A75</sub> (clone 11) was

determined to have the type A cDNA in the correct orientation with respect to the MMT promoter.

5) Plasmid DNA from recombinant pdMMTNeo  $\alpha$ -AE<sub>A75</sub>(clone 11) was digested with BamHI. The linearized vector was isolated and purified and then treated with bacterial alkaline phosphatase (B.A.P.) for 2 hours at 37°C to remove 5' phosphates. The BPV-1 genome was isolated and purified following B and BamHI digestion of the vector pdBPVMMTNeo. This BamHI fragment of BPV-1 DNA, which is approximately 8.0 kb, was then ligated to the BamHI linearized and B.A.P. treated pdMMTNeo  $\alpha$ -AE<sub>A75</sub> vector, for 3 hours at 14°C. After the ligation mixture was transformed into *E. coli* RR1, the recombinant clones were selected on 50 ug/ml Ampicillin LB agar plates. The recombinant plasmids were analyzed for BPV DNA and were also analyzed for type A AE cDNA. Restriction mapping revealed that clone 21 was approximately 17 kb and produced a restriction map as expected. This expression plasmid was then used for expression of  $\alpha$ -AE<sub>A75</sub> in mouse C127 cells.

6) Mouse C127 cells were transfected with 20 ug of pdBPV-MMTNeo  $\alpha$ -AE<sub>A75</sub> by the standard CaPO<sub>4</sub> precipitation technique. Approximately 2 weeks post transfection, transformed foci were individually picked and grown in growth media containing the antibiotic G418. When cells were grown to a sufficient capacity in Dulbecco's Modified Eagle Medium plus 10% fetal calf serum, the clones were assessed for the ability to secrete Alpha Amidating Enzyme by measuring the enzymatic activity in the conditioned cell culture media, as well as by measuring the alpha amidating enzyme immunoreactivity in the medium using standard radiolabelling and immunoprecipitation techniques. Clones secreting active, immunoreactive 75 kD alpha amidating enzyme were expanded to large numbers of cells (switched to cell culture medium with reduced serum and therefore reduced level of exogenous protein) and are in use to produce large quantities of active recombinant enzyme from the cell conditioned media.

The terms and descriptions used herein are embodiments set forth by way of illustration only, and are not intended as limitations on the many variations which those of skill in the art will recognize to be possible when practicing the present invention, as defined by the following claims.

#### Claims

1. A host capable of expressing the polypeptide sequence of an alpha amidating enzyme, said host comprising an expression vector which includes a transcriptional promoter followed downstream by a DNA sequence foreign to said host which encodes said amidating enzyme, said vector being capable of directing expression of polypeptides in said host.

2. A host capable of expressing the polypeptide sequence of an alpha amidating enzyme, said host comprising an expression vector containing a transcriptional promoter followed downstream by a DNA sequence foreign to said host which is capable of hybridizing under stringent conditions with a DNA sequence of:



[illegible]

3214

55

-355

5  
10  
15  
20  
25  
30

GGGATGCTGCCAGTGGGACCAGGTTCGAATGACTGCGCGGGTCTGCGCGATAGCCCTCACAGCCCTGGCATGGCAGCGATAAGGACCTGGCTCC -246  
TCCAGCCCTGTGGTCCCAAGCCCTCTGGCCCTCTCGAGCGGGGGGGCTCCAGCGGGGGGGAGGGGACTCCAGAGCCCTTCCACCGCTCCGA -146  
CCGACGAGAGCCCAAGCCCTTGGCGCGCATGCGCCCTGCGCGGGCATGAAGTACCTGCGCCAGCCCTGGCATCTGCGCGCGCGCGCAUCCCTGCCA -46  
TCCCGCGCGCGCGCATGCGCCCTGCGCTGCGCGGACATGCGCGCGCGCGCGCGCGGTCTGCTACTGCTGCTGCGCGCTGCTGCGCG -53

..... m s s r r s s g l l l l l l g l l a l (119)

TGCAGAGCACTGCGCTGCGCTTCAGAAGCCCACTTTCTGCTTTAAGAGCTTAAAGAACTACCAGATCATTTTCCAATGAATGCTTGGTACCATGG -155  
q s s c l a f r s p l a v f k r f k o t t r a f s n o c l g t l g

ACCACTCAGCCCTCTTGAATGCATCAGATTTTGGCTGGATATTCGATCCCTGGGTTACACCTAAGAGTCTGACACATACTTCTGCATGCCATGGT -255  
p v t p l d a s d f a l d i r s p g v t p h o s d t y f o s s r

CTGCGCTGCGATGACGAAGCCCTTCTGATGACTTCAAGCCCTGCGCCAGCATGATATCTCCACCATATGCTGCTGTTCCATGCAATATGCGCTGT -355  
l p v d o e s f v i d f k p r s a s d t y h h a l l i g c a p s s (119)

CCACTGGAAGTTACTGTTTTGATGAAGGAACCTGTACAGATAAAGCCCAATATTCTATATGCGTGGCAAGGAATGCTCCCGCCACCGGCTCCGAA -455  
t g s y w f c d o g t c t d k a n l l y a w a r m a p p t r i p k

AGGTGTTGGATTGAGAGTTGGAGGAGAACTGGAAGCAAAATCTGCTCTTCAAGTTCACTATGCGCATATCAGTCTTTTCCAGATAATCACAAGAC -555  
s v s f r v s s o t g s h y f v l q v h y g d l s a f r d s h k d

TGCTCTGCGTGTCCGTACATCTCACAGGTGTGCGCCAGCCCTTAATTCGCGGCATGTACCTTATCATGCTGTGTCACACTGTCATACCACGAGGAGA -655  
c s s v s v h l t r v p q p l l a g s y l a s s v d t v l p p g e k (219)

AAGTAGTAATCTGACATTTCTGCGCAATACAAAATGTATCCAATGCATGTGTTGCGTACAGAGTCCACACTCAUATTAGGTAAAGTGGTGAGCGG -755  
v v a a d i s c q y k n y p h v f a y r v h t b b i g h v v s s

ATACAGAGTAAGAAACGACAGTGGACACTGATTGGAGCCGACAGCCCGAGCTGCCACAGGCTTTCTACCCCTGCGAACACCCCGTTGATGTTACTTT -855  
y r v r n g q u t l i g r q a p q l p q o f y p v o h p v d v t f

GGTGATATAGTGGACGCAAGATGTGTCTCACTGGTGAAGCGACGACAGGCCACCCACATCGGCGCCACTTCTAGTGACGAAATGTGTAACTGTACA -955  
s d l i a a r c v f t g e g r t o a t h i g s t s d e n c a l y i (319)

TCAATGATTACATGGAAGCCAAATATGCACTTCTCTCATGACCTGTACAAAGAACGTGGCTCCAGATATGTTCAAGAACTATCCGAGCAGGCCAATAT -1055  
u y y n o a k y a l s f o t c t k n v o p d u f r t i p a e a n i

CCCAATTCCTGTCAAACCGGACATGTTATCATCCAGCGGCATCACAAGAGGAGGAGAAACAAAGAAAGAGTCTTTAATCCAGCCCAAAACAGGGA -1155  
p i p v k p d v s n h g h h k s s o k s a l s q q p k q s

CAGGAGAGATATTAGACGAGCTGATTTCTATTCACTGCTTTCCAGCTGCTACGAGAAAGGAGATGTTATGTGCACAAGTATAATCTTACAGAAA -1255  
o e v l o q g d f y a l i s k l i g e r o d v h v h k y a p t o k (419)

AGACACAATCTGGCTCAGACCTCGTACCTGACATTCCAACCTGGTCCAGAAAAGCACCTTCTGCTCTCAGCCAGAGAACTGCAGAGCATGAGGA 1355  
 t o o d l v o o i n a v v q k b d l s r d a r o o o o o  
 5 ATGGGGTAATCTATCTAOTCAAGACAGGATCCACAGATTCCACCACTAGAGTCAACTCTGAGGCCAGCTCAGACAGACCTTTCTCTCCACCAO 1455  
 u g n a l l v r d r i h r f h q i o o t i r p o o s r a f i e q q  
 CCTGCCAAGCCCTTGGCAACCACAACCTCAGGAGATTTCATGTGGAAGAAGAACTGGACTGGCTGGAGTGTACTTGTACCAGGCCAGCTTTCTG 1555  
 p o o s p u o p p a s d i h v o o l d u p e v y l l p s q v s (519)  
 CGGTGGCCCTGGATTCTAAGAATAACCTAGTGAATTTCCACAGAGGTGACCATGTTTCGGATCGAAACTCTTTTACAGCAAGTTTGTACCACAAAG 1655  
 v a l d s h a n l v l f b r g d h v u d s n e f d a k e v y q q r  
 10 AGCTTTGGCCCAATTGAAGAAGACACCATCTGCTCATTGACCCAAATATGCTGAAATCTCCAGTCCAGTGGCAAGACCTGTTTATTACACAC 1755  
 g l s p l o o d t l l v l d p n n a l l q o s g k l l f y l p h  
 GGCTTCAGCATAGATACAGATGGAATTTATGGCTCAGATGTGGCTCTCCACCAAGTGTTCAAATTCGACCCCATACCAAGAAGGCCCTCTGTAA 1855  
 s l s i d t d g n y u v t e v a l h q v f k l d p h s h o g p l l l (619)  
 15 TTCTGGCAAGCAGCATGCAACCTGGAGTGAACAAATCTTTCGCCAGCCACCGATGTGCTGTGAGCCCACTACTGCAGCTGCTCTCTGTGACA 1955  
 l s r s o q p p s d o n b f c p t l d v a v o p s l g o v l v a d  
 CGGTACTGTAAACAGTCCGATTGTGACGTTTCCACCAAGCCGAAAGTTCCTCAGCCAGTGGGAGAGAGTCTCTGCAAGCACTCTAGCCAGGCCAG 2055  
 s y c a s r l v q f s p s k l v t q u g o o o s s a p r p s q  
 20 TTCAGTGTCTCTACAGTTTGGCCCTTGTGCTCATTTGGACCACTGTGTGTGGCAGACAGGGAATTCGCGCAATCCAATCTTCAAACTGACACCA 2155  
 f s v p h a l a l v p h l d e l c v d r o n s r l q c f h t d t h (719)  
 AAGAATTTGTGACAGCAGATTAAACCAAGCATCTTGAAGCAATGTCTTTCGCAATTCATATATACCAAGTTTCTCTTTCGCGTAACGGGAAGCCCTTA 2255  
 o l v r o l k h a s f g r n v f a l a y l p s f l a v n g k p y  
 CTTTGCAGACCAAGAGCCCTGCAAGGATTGTGATGAATTTCCAGTGGGGAATTATAGACGCTCTCAAGCCAGTACCAAGCACTTCCACATGCGT 2355  
 f s d q o p v q s f v n d s s o l i d v t k p v r k h l d p  
 25 CATGATATTTGGCTTCTGAAGATGGCACTGTGTACATTGACAGCCACACAAACACCTGTGGAAGTTCACCTCACTGAAAAATGGAGCATCGGT 2455  
 h d l v o s o e g t v y l s d o h t n t v n k f t l t o k o b r s (819)  
 CAGTCAAAAGCCTGGCATTGAAGTCCAGGAATCAAAAGCCGAGGAGTGTTCACCAAGTTCGAGAACAAACCCACCTCTCAGAAATTCAGCAAGAT 2555  
 v h k a g l o v q o l k a o o v v o p k v o n h p t a o o l q k o  
 30 GCAAGACAACAGAACTGACCAAGAGCCCGCTCCGAGTGTCCGTGCTCATTACAACCCCTCTGCTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCT 2655  
 q o k q k l s t o p s s v s v l l t t l l v l p v l v l l a l  
 CTCATGTTTATTCGTTGCAAAAATCAAGGCCCTTGGACCAAGATCATGCCCAAGCTCGAGTCAACTTCTGCAAGAGTCTGGGAAGAAGAGGAGCAG 2755  
 v s f i r u h k s r o f g o d h d r k l o n o o g r v l s r s s (919)  
 35 GAGGAGCAACAGAGAG 2855  
 s s r o e s t e r o s s r h o o s s r f o s s r f (944)  
 ACCGCTCTCACAATAAGTACAAGCGGAAGCGGATTCCTTAAGCACCGATTCAAGCTGTCTAGTCTTTCGCTGAGATGCAATGTGCT 2945

4. A host according to any preceding claim, wherein the DNA sequence includes a stop codon upstream from a sequence which would otherwise code for a membrane spanning region.

5. A host according to claim 4, wherein the stop codon is upstream from a sequence which would otherwise code for consecutive lysine residues.

6. An expression vector capable of directing the expression of a polypeptide sequence of an alpha amidating enzyme in prokaryotic cells, said vector comprising a transcriptional promoter followed downstream by a first DNA sequence having an amidating enzyme-coding region, said first sequence being sufficiently homologous to a natural DNA sequence for expressing natural amidating enzyme to undergo hybridisation with said natural sequence under stringent conditions, and said first sequence including an initiating methionine codon within about 50 nucleotides of the start of said enzyme-coding region.

7. An expression vector according to claim 6, wherein the first three codons of the enzyme-coding region encode the amino-acid sequence Ser-Phe-Ser or Ser-Leu-Ser.

8. An expression vector according to claim 7, wherein the initiating methionine codon immediately precedes the DNA sequence encoding Ser-Phe-Ser or Ser-Leu-Ser.

9. An expression vector according to any of claims 6 to 8, wherein the initiating methionine codon is within about 50 nucleotides upstream from the start of the enzyme-coding region.

10. An expression vector according to any of claims 6 to 9, wherein the enzyme-expressing sequence includes a stop codon upstream from a sequence which would otherwise code for a membrane-spanning domain.

11. An expression vector according to claim 10, which has no internal initiation site between the initiating methionine codon and the stop codon.

12. An expression vector according to claim 11, wherein all DNA sequences coding for ribosome binding sites which are located between the initiating methionine codon and the stop codon are located sufficiently distant from an initiating codon to prevent internal initiation and translation of a polypeptide in quantities observable as a separate band after electrophoresis on SDS-polyacrylamide gel.

13. An expression vector capable of directing the expression of a polypeptide sequence of an alpha amidating enzyme in eukaryotic cells, said vector comprising transcriptional promoter followed downstream by a first DNA sequence having an amidating enzyme-coding region, said first sequence being sufficiently homologous to a natural DNA sequence for expressing natural amidating enzyme to undergo hybridisation with said natural sequence under stringent conditions; and said first sequence including a stop codon upstream from a sequence which would otherwise code for a membrane-spanning domain.

14. An expression vector according to any of claims 6 to 13, wherein the natural DNA sequence is as defined in claim 2 or claim 3.

15. 15. An expression vector according to claim 14, wherein either the natural DNA sequence is as defined in claim 2 and the stop codon is located between bases 2025 and 2275, or the natural DNA sequence is as defined in claim 3 and the stop codon is located between bases 1148 and 1492 or between bases 2340 and 2690.

16. An expression vector according to claim 14, wherein the natural DNA sequence is as defined in claim 2 and the stop codon is located between bases 1148 and 1208.

17. An expression vector according to any of claims 6 to 16, wherein the stop codon is within five codons of a DNA sequence coding for consecutive basic amino-acid residues.

18. An expression vector according to claim 13, wherein the natural DNA sequence is as defined in claim 3, and which is obtainable by ligating said first DNA sequence into an expression system selected from pDBPVMMTNeo, pSV2, pRSV, pMAMNeo, peuK-C1, PCH110, and derivatives thereof.

19. A prokaryotic host comprising an expression vector according to any of claims 6 to 12 or to any of claims 14 to 17 when appendant to any of claims 6 to 12.

20. A eukaryotic host comprising an expression vector according to claim 13 or claim 18 or to any of claims 14 to 17 when appendant to claim 13.

21. A eukaryotic host according to claim 20, which is selected from IVI deposit 1029, Hela, CV1, C127, CHO (Chinese Hamster Ovary) and COS.

22. A host capable of expressing the polypeptide sequence of an alpha amidating enzyme, said host comprising an expression vector containing a transcriptional promoter followed downstream by a DNA sequence as defined in claim 2 or claim 3.

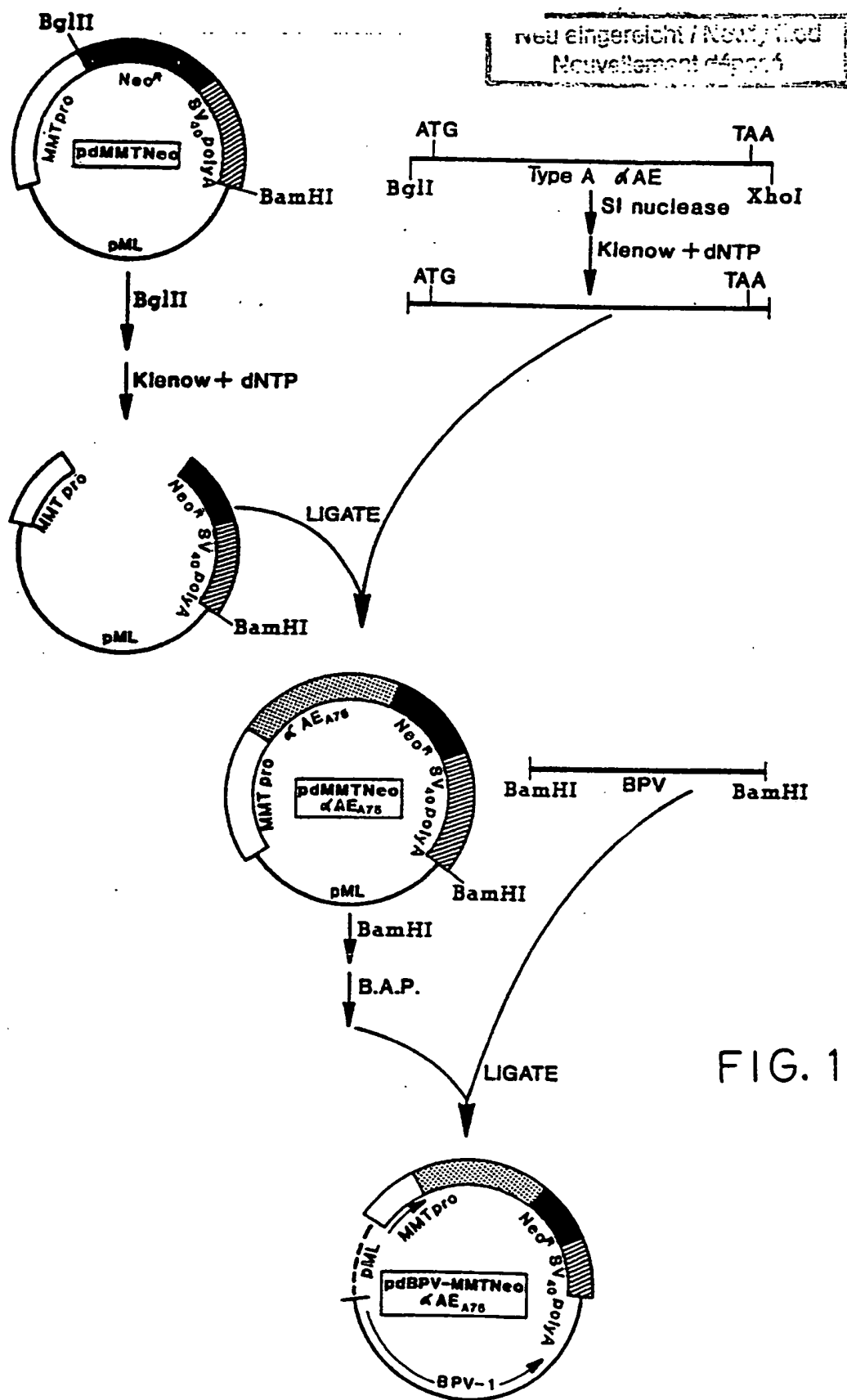


FIG. 1

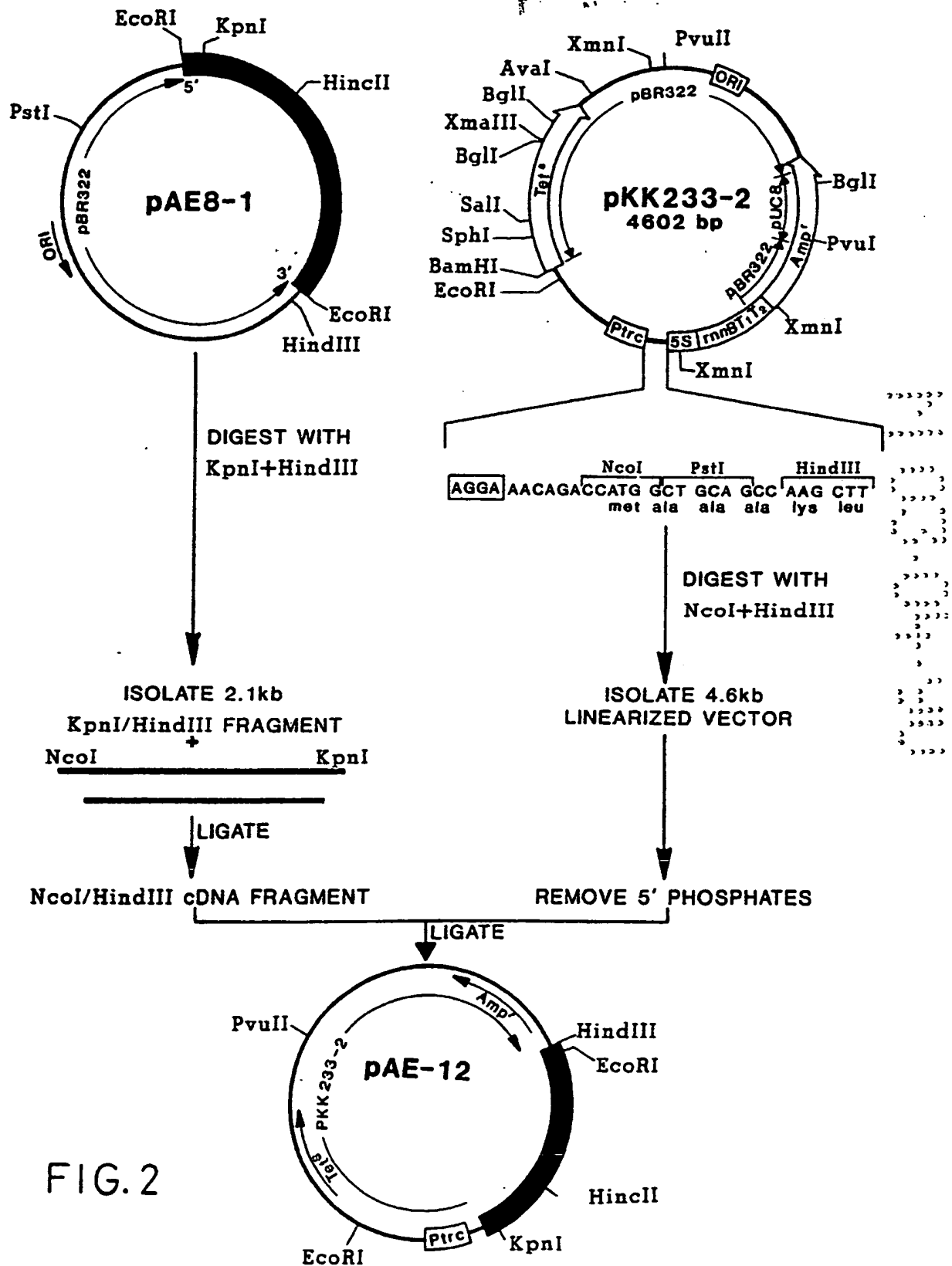


FIG. 2

Not a registered trademark  
 of the American Chemical Society

FIG. 4  
 C  
 PAE 51  
 PAE 36  
 PAE 31  
 PAE 24  
 PAE 12  
 B 1 2 3 4 5 6 7 8 9 10 11 12  
 + - + - + - + - + - + -

118  
 68  
 43  
 29

FIG. 3  
 C  
 PAE 51  
 PAE 36  
 PAE 31  
 PAE 24  
 PAE 12  
 A 1 2 3 4 5 6 7 8 9 10 11 12  
 + - + - + - + - + - + -

118  
 68  
 43  
 29

118  
 68  
 43  
 29



(19)



Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11) Publication number:

**0 382 403 A3**

(12)

## EUROPEAN PATENT APPLICATION

(21) Application number: 90301034.6

(51) Int. Cl.<sup>5</sup>: **C12N 15/53, C12P 21/02**

(22) Date of filing: **01.02.90**

(30) Priority: **06.02.89 US 307366**

(43) Date of publication of application:  
**16.08.90 Bulletin 90/33**

(84) Designated Contracting States:  
**AT BE CH DE DK ES FR GB GR IT LI LU NL SE**

(98) Date of deferred publication of the search report:  
**10.04.91 Bulletin 91/15**

(71) Applicant: **UNIGENE LABORATORIES INC.**  
**110 Little Falls Road**  
**Fairfield, NJ 07006(US)**

(72) Inventor: **Betelsen, Arthur H.**  
**599 Maitland Avenue**  
**Teaneck, New Jersey(US)**  
Inventor: **Mehta, Nozer M.**  
**8 Apache Trail**  
**Rockaway, New Jersey(US)**  
Inventor: **Beaudry, Gary Agide**  
**81 Inwood Avenue**  
**Montclair, New Jersey(US)**

(74) Representative: **Perry, Robert Edward**  
**GILL JENNINGS & EVERY 53-64 Chancery**  
**Lane**  
**London WC2A 1HN(GB)**

(54) Expression systems for amidating enzyme.

(57) Alpha-amidating enzyme is produced by recombinant DNA techniques recoverable in high yields and at high purity. Both eukaryotic and prokaryotic expression vectors are provided having a transcriptional promoter followed downstream by a DNA sequence which encodes amidating enzyme. The vector selected is one capable of directing the expression of polypeptides in the host selected, and preferred hosts are transected with the described vectors.

**EP 0 382 403 A3**



European Patent  
Office

# EUROPEAN SEARCH REPORT

Application Number

EP 90 30 1034

| DOCUMENTS CONSIDERED TO BE RELEVANT   |  |  |   |
|---|--|--|---|
| Category  | Citation of document with indication, where appropriate, of relevant passages  | Relevant to claim                              | CLASSIFICATION OF THE APPLICATION (Int. Cl.5)                               |
| X<br>A  | EP-A-0 299 790 (SUNTORY LTD. et al.)<br>* whole document *   | 1,6-9,<br>19<br>2,3,13,<br>20,22               | C 12 N 15/53<br>C 12 P 21/02  |
| X   | MOLECULAR ENDOCRINOLOGY vol. 1, no. 1, 1987, pages 777-790, US; B.A. EIPPER et al.: "Structure of the Precursor to an Enzyme Mediating COOH-Terminal Amidation in Peptide Biosynthesis"<br>* whole article *   | 1-3,6,<br>19,22                                |   |
| P,X   | MOLECULAR ENDOCRINOLOGY vol. 4, no. 1, 1990, pages 132-139, US; S.N. PERKINS et al.: "Stable Expression of Full-Length and Truncated Bovine Peptidylglycine alpha-Amidating Monooxygenase Complementary DNAs in Cultured Cells"<br>* whole article * | 1-4,13,<br>14,20,<br>22                        |   |
| P,X   | WO-A-8 902 460 (GENENTECH INC. et al.)<br>* whole document; in particular page 7, line 13 - page 8, line 30 *  | 1-3,6,<br>13,20,<br>22                         | TECHNICAL FIELDS<br>SEARCHED (Int. Cl.5)<br><br>C 12 N 15/53<br>C 12 N 9/02 |
| D,P<br>X  | EP-A-0 308 067 (UNIGENE LABORATORIES INC.)<br>* abstract; page 6, lines 4-14; page 7, line 54 - page 8, line 4; page 17 example 7; page 26, example 10; claims 9,10 *  | 1,6,13,<br>20,22                               |   |
| The present search report has been drawn up for all claims  |  |  |   |
| Place of search<br>BERLIN   |  | Date of completion of the search<br>14-11-1990 | Examiner<br>JULIA P.  |
| <b>CATEGORY OF CITED DOCUMENTS</b><br>X : particularly relevant if taken alone<br>Y : particularly relevant if combined with another document of the same category<br>A : technological background<br>O : non-written disclosure<br>P : intermediate document<br><br>T : theory or principle underlying the invention<br>E : earlier patent document, but published on, or after the filing date<br>D : document cited in the application<br>L : document cited for other reasons<br><br>A : member of the same patent family, corresponding document |  |  |   |



European Patent  
Office

## EUROPEAN SEARCH REPORT

Page 2

Application Number

EP 90 30 1034

| DOCUMENTS CONSIDERED TO BE RELEVANT  |  |  |   |
|--|--|--|---|
| Category   | Citation of document with indication, where appropriate, of relevant passages  | Relevant to claim                              | CLASSIFICATION OF THE APPLICATION (Int. Cl.5) |
| X  | PROC. NATL. ACAD. SCI. USA vol. 86, no. 2, January 1989, pages 735-739, Washington D.C., US; D.A. STOFFERS et al.: "Alternative mRNA splicing generates multiple forms of peptidyl-glycine alpha-amidating monooxygenase in rat atrium"<br>* whole article; in particular page 736, figure 1; page 738, figure 3 * | 1-3,6, 19,22                                   |   |
| A  | WO-A-8 602 099 (UNIGENE LABORATORIES INC.)<br>* whole document *   | 1  |   |
| E  | WO-A-9 008 190 (SUNTORY LTD.)<br>* whole document, in particular page 16, line 5 - page 23, line 23 *  | 1,6,13, 19,20                                  |   |
| A  | EP-A-0 249 477 (IMMUNEX CORP.)<br>* whole document, in particular claims 1,6,8-10 *  | 4,10,13  |   |
| T  | BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS vol. 169, no. 2, 15 June 1990, pages 551-558, New York, US; J. GLAUDER et al.: "Human Peptidylglycine alpha-Amidating Monooxygenase: cDNA, cloning and Functional Expression of a truncated form in COS s" cells" * whole article *                            | 1-3,13, 20,22                                  |   |
| The present search report has been drawn up for all claims   |  |  |   |
| Place of search<br>BERLIN  |  | Date of completion of the search<br>14-11-1990 | Examiner<br>JULIA P.                          |
| <b>CATEGORY OF CITED DOCUMENTS</b>   |  |  |   |
| X : particularly relevant if taken alone<br>Y : particularly relevant if combined with another document of the same category<br>A : technological background<br>O : non-written disclosure<br>P : intermediate document  |  |  |   |
| T : theory or principle underlying the invention<br>E : earlier patent document, but published on, or after the filing date<br>D : document cited in the application<br>L : document cited for other reasons<br>A : member of the same patent family, corresponding document |  |  |   |

EPO FORM 1503 (01.91) (P0001)